



Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

On making public policy with publicly available data: The case of U.S. communications policymaking[☆]

Philip M. Napoli^{a,*}, Joe Karaganis^b^a Fordham University, United States^b Social Science Research Council, United States

ARTICLE INFO

Available online xxxx

Keywords:

Open government

Data access

Communications policy

ABSTRACT

A fundamental principle of public policymaking should be that public policy must be made with publicly available data. This article develops this position and applies it to an assessment of the current state of communications policymaking, a policy area in which controversies surrounding the transparency of policy research and the accessibility of policy-relevant data have been both common and extremely contentious in recent years. This article provides a detailed assessment of the challenges confronting greater transparency and accessibility of communications policy-relevant data, as well as a detailed set of proposals for improving the current situation, in an effort to build towards an environment in which public policy is made with publicly available data.

© 2010 Elsevier Inc. All rights reserved.

Introduction

It has been well-documented how research has come to play an increasingly influential role in public policymaking (see, e.g., Heinrich, 2007; Sanderson, 2002; Warren, 2002). Government agencies not only conduct their own research and analyses to guide their decision-making, but also solicit, and rely upon, research conducted and submitted by third parties such as industry organizations, public interest advocacy groups, and scholars. Today, "Information is the lifeblood of regulatory policy" (Coglianese, Zeckhauser and Parson, 2004, p. 277).

Given the increasingly influential role that research plays in policymaking, it is important that we closely scrutinize how policy research is conducted and utilized (see, e.g., Wagner, 2009; Warren, 2002). The integrity and effectiveness of contemporary policymaking is very much a function of the integrity, accuracy, and objectivity of the research that guides policy decision-making. In today's policy-making environment, in which an increased emphasis is being placed on the transparency of government decision-making (see, e.g., Knight Commission on the Information Needs of Communities in a Democracy, 2009; Open government: A progress report to the American people, 2009; Robinson, Yu, Zeller and Felton, 2008/

2009), it is vitally important that we consider the transparency surrounding the research that guides government decision-making.

This article focuses on the question of whether adequate mechanisms are in place for ensuring the needed levels of transparency and data accessibility for the research that increasingly guides policy decision-making. The integrity of the policymaking process depends on there being adequate levels of transparency and accessibility for the data that guide these decisions. Thus, a fundamental premise of this analysis is that public policy should be made with publicly available data.

This article uses U.S. communications policymaking as a focal point for analyzing the core issues surrounding the transparency and accessibility of policy-relevant research and data. Recent years have seen a flurry of controversies surrounding research transparency and data access in the communications policy sector, revolving around high profile issues such as broadband deployment, media ownership, and localism (see, e.g., Napoli, 2008). In recent years, government agencies such as the Federal Communications Commission (FCC)¹, the National Telecommunications and Information Administration (NTIA)², and the U.S. Government Accountability Office (GAO)³ have initiated proceedings and investigations dealing specifically with issues related to data quality, transparency, and access in the communications sector. As this growing level of activity suggests, it has become clear that systematic

[☆] This research was an outgrowth of the Data Consortium initiative of the Social Science Research Council's Necessary Knowledge for a Democratic Sphere Program, which was supported by the Ford Foundation's Electronic Media Policy Portfolio.

* Corresponding author.

E-mail addresses: pnapoli@fordham.edu (P.M. Napoli), karaganis@ssrc.org (J. Karaganis).

¹ See, for example, <http://reboot.fcc.gov/data/>. Data transparency and access have become priorities in the new administration's reorganization and reconfiguring of the FCC.

² See, for example, the NTIA's recent Broadband Data Transparency Workshop (http://www.ntia.doc.gov/advisory/broadbanddata/WorkshopTranscript_10302009.txt).

³ See, for example, U.S. Government Accountability Office (2006, 2009).

policy changes are necessary in this sector in order for research to serve policymaking in ways that meet democratic standards of transparency and accountability.

This article addresses this issue via an analysis of a range of primary sources, including government agency decisions, court filings and decisions, congressional hearings and testimony, government reports, and stakeholder filings with government agencies. The first section of this article outlines the fundamental premise that public policy should be made with publicly available data. The second section applies this standard to the current state of U.S. communications policymaking and identifies significant shortcomings in existing practices and legislative safeguards. The third section offers a set of proposals for addressing these shortcomings and for creating a communications policymaking environment that better approaches the ideal in which public policy is made with publicly available data. The concluding section considers the implications of this analysis of communications policy to public policymaking more broadly.

The case for making public policy with publicly available data

The issue of access to the research and data that drive policy decision-making strikes at the core of the functioning of a representative democracy and the role of information in the democratic process. Effective democracy depends upon an informed citizenry; in particular, it depends upon a citizenry that is well-informed about the activities of its government and the reasons that the government acts in particular ways (Cohen & Hahn, 1999; Knight Commission on the Information Needs of Communities in a Democracy, 2009).

As was noted above, research has come to play an increasingly influential role in government decision-making. Any piece of research, however, represents a complex input into the policymaking process. A piece of research, and the conclusions contained within it, are the result of a series of subjective decisions at each stage in the research process, from conceptualization to methodological design to the data-gathering process to the statistical analyses employed. As influential inputs into policymaking, these decisions need to be subjected to the same kind of scrutiny and analysis that are applied to other aspects of the policymaking process. The strengths, weaknesses, and potential biases of policy research need to be subjected to the strictest possible scrutiny, which can only be achieved if all elements of the research, ranging from methodological details to the underlying data, are transparent and accessible.

Legitimate concerns regarding the public's confidence in its policymakers arise from any policymaking process that relies upon data and analysis that cannot be subjected fully to public scrutiny and reassessment; thus, "[p]ublic access to data ensures greater transparency, which lends legitimacy to the regulatory process" (Cohen & Hahn, 1999, p. 536). To the extent that the inability to access policy-relevant data undermines this transparency, citizens' confidence in their government's decision-makers suffers.

Moreover, a truly *representative* policymaking process can only be achieved via the accessibility of policy-relevant data. The greater the diversity of stakeholders that have the ability to meaningfully participate in the policymaking process—for example, by conducting analyses that feed into this process—the greater the likelihood that the information that guides decision-makers will reflect the full range of policy options and concerns. This kind of diversity of participation—and the public benefits it provides—can only be achieved if access to policy-relevant data is widespread. Ultimately, as the National Research Council (2005) has noted, "The benefits of providing wider access to microdata for researchers and policy analysts are better informed public policies" (p. 1).

These concerns have resulted in some legislative efforts to enhance access to policy-relevant data. The *Data Access Act* (1998), for instance, requires that the data generated by federally-funded studies be made available to requesting parties through the Freedom of

Information Act. The *Data Quality Act* (2000) directs the Office of Management and Budget to issue guidelines to federal agencies for "ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information)" that they disseminate (§ 515(a)). The Act mandates that these guidelines "establish administrative mechanisms allowing affected persons to seek and obtain correction of information maintained and disseminated by the agency that does not comply with the guidelines issued under subsection (a)" (*Data Quality Act*, 2000, § 515(b)(2)(B)). It is within this language that the *Data Quality Act* also provides a mechanism for public access to policy-relevant data. And, of course, the Freedom of Information Act provides an avenue via which the public can gain access to policy-relevant data. However, as this article will illustrate, the effectiveness of these legislative safeguards is limited in important ways (see below).

Also, it is important to recognize that concerns about public access to policy-relevant data apply not only to the research and data generated by government agencies, but also to the research and data generated by third parties in their formal filings with government agencies. Government agencies have come to rely extensively on such third-party research in their decision-making, as these agencies' commitments to internally generated data gathering and research have diminished in conjunction with the decades-long deregulatory trend (Wagner & Michaels, 2004). And, of course, such research typically supports advocacy for very specific policy positions. Such research therefore needs to have its underlying assumptions, data, and methodological approaches subjected to independent verification and re-analysis, whether by the government agencies themselves or other interested stakeholders. Otherwise, flawed or biased research and data can have an inappropriate degree of influence over policy decision-making (see, e.g., Wagner & Michaels, 2004; Warren, 2002).

This need for access and transparency surrounding third-party research is compounded when we consider the resource and participation imbalances that characterize today's increasingly research-driven policymaking process. To the extent that certain categories of stakeholders (e.g., industry groups) often have far greater resources than other categories of stakeholders (e.g., public interest advocates, scholars) for purchasing and gathering the data necessary for conducting and submitting research, the policy process is plagued by "information asymmetries" (Napoli & Seaton, 2007) that can allow certain stakeholders to have a disproportionate influence over policy outcomes. Such concerns reflect the fact that "Scientific knowledge, like land, labor, and capital, is a resource—indeed a commodity—and the ability to manipulate and control this resource has profound implications for the distribution of political power in democratic societies" (Nelkin, 1979, p. 118).

In the end, from a purely normative perspective, it seems fairly clear that in a well-functioning democracy public policy should be made with publicly available data. Increasingly privileged and unequal access to the raw data that guide policy decisions represents a significant failing in the construction of our policymaking process and, consequently, presents a significant roadblock to effective public policymaking and public confidence in policy decisions. Flawed data potentially lead to questionable policy outcomes; incomplete data potentially lead to questionable policy outcomes; biased research potentially leads to questionable policy outcomes; and influential research inputs that overwhelmingly originate from certain categories of stakeholders to the neglect of others potentially lead to questionable policy outcomes. These are the dangers that transparency of policy research and accessibility of policy-relevant data can combat.

The case of U.S. communications policymaking

Concerns about data transparency and access have become particularly pronounced in the communications policy sector in

recent years. As one recent report noted, “the communications policy debate in the United States today is inseparable from debates about the data used to make claims about policy propositions” (Flamm, Friedlander, Horrigan and Lehr, 2007, November, p. 2). Recent years have seen an array of controversies emerge around the role of research and data in communications policymaking, including accusations of biased and suppressed research, claims of biases in the selection of researchers to conduct government-commissioned studies, and constant criticisms of the thoroughness and reliability of the data gathered by communications policymakers on a variety of fronts (see Napoli, 2008).

These issues have attracted the attention of Congress, not only in the form of congressional hearings (e.g., *U.S. House of Representatives*, 2007), but also in the form of an ongoing U.S. Government Accountability Office investigation that is specifically examining the Federal Communications Commission’s data collection practices.⁴ A complete account of the wide-ranging contexts in which issues of data access and transparency have come to the fore of communications policy debates is beyond the scope of this section (for examples of the growing body of work addressing this increasingly prominent topic, see Cramer, 2009; Frieden, 2009; Napoli, 2008; Napoli & Seaton, 2007; Sandoval, in press). Rather, this section focuses on distilling the key data transparency and access obstacles that connect the various controversies that have arisen across a number of disparate communications policy contexts.

Commercially sensitive data

One significant obstacle to access to policy-relevant data involves concerns over the commercially sensitive nature of the data that are sometimes obtained by policymakers through either voluntary or mandatory data solicitations from regulated firms. Effective policy analysis often requires the gathering and analysis of data that potentially can be considered commercially sensitive. Such situations require a careful balancing of the economic concerns of the regulated firms with the broader public interest concerns associated with transparent and well-informed policymaking (Anderson & Seltzer, 2009).

Such concerns are at the core of the ongoing debate over public access to the broadband deployment data being gathered by the Federal Communications Commission and the National Telecommunications and Information Administration (NTIA) (see, e.g., *Federal Communications*, 2010). Rapid broadband deployment has become an increasingly important communications policy goal in recent years, and the FCC has come under substantial scrutiny for releasing assessments of broadband deployment based on incomplete and flawed data, and according to many accounts, for dramatically overstating the extent of broadband availability in the U.S. (Flamm et al., 2007; Ford, 2007, May 17; *U.S. Government Accountability Office*, 2006, 2009). Critics have gone so far as to accuse the FCC of employing a flawed methodology in data gathering and analysis in order to obscure the relatively poor state of broadband penetration in the U.S. relative to other developed nations (see, e.g., Frieden, 2009). The result has been legislation directed at improving the collection of broadband penetration data (*Broadband Data Improvement Act*, 2008).

Throughout this process, the FCC has emphasized the need for wide access to these data, noting that “public availability allows consumers and experts the opportunity to review the data to ensure the accuracy of the information” (*Federal Communications Commission*, 2000, para. 86). Further, the Commission stated,

by allowing public release of as much of the information as

possible, associations, scholars, and others will be able to use the information in their independent analyses of Commission policies, thereby aiding the Commission in crafting regulations that address specific market problems and eliminating those regulations that have outlived their usefulness (*Federal Communications Commission*, 2000, para. 96).

Despite these statements, which perfectly encapsulate why open data policies serve the public interest, the Commission has resisted making the detailed broadband penetration data publicly available.

The Commission has denied Freedom of Information Act (FOIA) requests for the data, based on the conclusion that some of the information contained within the database falls within FOIA exemptions regarding commercially sensitive information (Burgee, 2006). Legal actions have failed to result in the FCC releasing the data, due to this “trade secrets” exemption (*Center for Public Integrity v. Federal Communications Commission*, 2007), which has been interpreted by the courts as being met if the disclosure of the information is likely to either impair the government’s ability to obtain necessary information in the future or to cause substantial harm to the competitive position of the person from whom the information was obtained (see *National Parks and Conservation Association v. Morton*, 1974).

In theory, these should be relatively high thresholds to meet, particularly in the context of broadband data gathering, where data provision has been deemed mandatory by the FCC (see Cramer, 2009) and where the language of the court’s interpretation of the trade secrets exemption emphasizes “substantial harm.” Thus, it would appear that if only *some* harm is likely to befall the providers of the information, then the trade secrets exemption has not been met. Rather, the likelihood of *substantial* harm must be demonstrated. Similarly, it is important to emphasize that there must be a *likelihood* of substantial harm. Thus, unlikely scenarios in which harm *might* occur do not meet the threshold established in the legal interpretation of the Act.

Critics continue to emphasize the significant gap between policymakers’ rhetoric of transparency and access pertaining to broadband penetration data and the reality in terms of stakeholder access to the data (see, e.g., Turner, 2009). This tendency persists to this day. Consider, for instance, that the NTIA’s October, 2009 Broadband Data Transparency Workshop dealt with a wide range of topics related to what *types* of data should be collected, but not at all with the issue of the transparency or accessibility of the data that are collected, despite the fairly clear mandate suggested by the event’s title. The issues of broadband data access were essentially tabled for an indeterminate later date.⁵

Further, in October of 2009, the FCC issued a protective order giving participants in the Commission’s broadband data-gathering activities the discretion to designate the materials they provide to the Commission as Highly Confidential (*Federal Communications Commission*, 2009). Highly Confidential information is defined as “information . . . that is not otherwise available from public sources and that consists of detailed or granular information regarding the location, type, or cost of last-mile infrastructure used by a Submitting Party to offer broadband services” (*Federal Communications Commission*, 2009, pp. 2–3). Such information will be only accessible to “Outside Counsel of Record, their employees, and Outside Consultants and experts whom they retain to assist them” in the proceeding (*Federal Communications Commission*, 2009, p. 3). Even amongst

⁴ The results of this investigation have not yet been released. The first author provided testimony to the U.S. GAO in connection with this investigation.

⁵ As one NTIA official stated in regards to the issue of data disclosure, “We are not going to solve those problems today, and let’s not get distracted by those.” This perspective was reflected in the statement of another NTIA official, who suggested rather vaguely that “I’m sure we will [address the disclosure issue] in other forums at some point or another.” These statements come from the workshop transcript available at: http://www.ntia.doc.gov/advisory/broadbanddata/WorkshopTranscript_10302009.txt.

those potentially able to gain access to this information, the terms of access are quite restrictive (Federal Communications Commission, 2009, p. 4).

The Federal Communications Commission (2000) has asserted that maintaining the confidentiality of the data that telecommunications operators file in regards to broadband deployment is necessary in order to maximize the voluntary compliance with the information collection. However, in the same document, the Commission describes participation in the broadband data collection as “mandatory” (Federal Communications Commission, 2000, para 91). Certainly, the issue of whether broader data access would affect the extent to which regulated firms cooperate with federal data-gathering efforts is a legitimate concern (see below). Today, however, in terms of maintaining the delicate balance between commercial interests and the public interest, the public interest is not being given sufficient weight.

What is frustrating in this particular case is that, just as the broadband penetration data being gathered by the Commission are being improved in many of the ways that were widely considered essential for the conduct of meaningful and insightful policy analysis (see, e.g., U.S. Government Accountability Office, 2009), those data are becoming virtually inaccessible to almost anyone wishing to conduct such analyses. If improvements in data quality are accompanied by diminishments in data access, it is questionable whether the overall public interest is in fact being better served.

Internal deliberations

Another common impediment to policy-relevant data involves the FOIA exemption which protects the confidentiality of regulators' internal deliberations. The deliberative process provision of the Freedom of Information Act is designed to protect the internal deliberations of government agency officials. The underlying rationales for the provision are to: a) encourage the free flow of ideas and the uninhibited exchange of views among government decision-makers; and b) prevent premature disclosure of proposed policies and thereby prevent the public from confusing policy options that are merely under discussion with those that being implemented (see Jensen, 1999).

This exemption was recently at the forefront of another data controversy in the Federal Communications Commission's 2004 broadcast localism proceeding. Georgetown University's Institute for Public Representation (IPR) submitted a FOIA request for all studies and/or proposals for studies, reports, analytical assessments, and factual data gathered or compiled by the FCC or by outsiders under contract with the FCC, that were related to the proceeding (*Institute for Public Representation v. Federal Communications Commission*, 2007). In response, the FCC made available a range of draft documents, supporting research, solicitations, proposals, and contracts, as well as sound recordings and spreadsheets (Perko, 2007). However, approximately 1400 pages of internal Commission records were withheld on the grounds that they not only contained confidential commercial information (including copyrighted materials), but also reflected the agency's internal deliberative process (Perko, 2007).

While it is reasonable that certain aspects of any government agency's deliberative process might need to remain confidential, it is more difficult to justify such treatment for the research—and particularly the data—that feed into such deliberations. Indeed, the inclusion of factual material such as data within the protective sphere of the deliberative process exemption has proven controversial, as the linkage between factual information and the underlying rationales for the internal deliberations exemption to FOIA seems tenuous at best (Jensen, 1999). The internal deliberations of agency officials should be able to be effectively protected without denying the public the ability to access and re-analyze the data that feed into governmental decision-making.

Restrictive data licenses

The third primary obstacle that has confronted expanded public access to policy-relevant data involves the nature of the access provisions in the contracts that government policymakers enter into with various commercial data providers. A significant, though largely neglected, story in the history of U.S. public policymaking and in the history of communications policymaking in particular, is the gradual process in which commercial data providers have adopted many of the data-gathering responsibilities that once were the province of government agencies (see, e.g., Napoli & Seaton, 2007). As a result, policymakers today often contract with various commercial providers for much of the data that they rely upon in their own analyses. These contracts contain strict access limitations that prevent external stakeholders—and even some individuals within the government agencies—from accessing the data (Napoli & Seaton, 2007).

This problem arose within the context of the FCC's localism proceeding as described above. In that proceeding, seven hundred pages of material were identified as spreadsheets and memoranda derived from proprietary data sources such as BIA Research and Warren Communications, for which access was restricted by the underlying commercial licenses (Perko, 2007).

Licensing restrictions also have been a recurring problem within the context of the FCC's quadrennial media ownership proceedings. For instance, in connection with its 2006 proceeding, the FCC commissioned a study titled *Ownership Structure and Robustness of Media*, which developed what was described as “one of the largest datasets ever assembled concerning ownership and the media” (Duwadi et al., 2007, p. 3). However, licensing restrictions on some of the data limited public access to only “a smaller set of data at both the DMA and station, system, and newspaper levels” (Duwadi et al., 2007 p. 3). Many of the central variables created for the database were withheld, including outlet revenues, outlet owner revenues, parent company revenues, audience shares, and minority ownership (Duwadi et al., 2007, pp. 20–23).

In these examples, because policymakers have come to rely heavily upon commercial data sources in their analysis, this reliance essentially creates a barrier to the necessary level of public access to the data that feed into the research that policymakers are conducting or commissioning to inform their decision-making. It is important to emphasize that the access provisions associated with legislation such as the Freedom of Information Act and the Data Access Act have been interpreted as not being applicable in these contexts, because these data are not deemed to be under the “control” of the government agency, given the restrictive provisions under which the agency has access to the data (see Napoli & Seaton, 2007).

In the end, when we consider the substantial controversies that have arisen in recent years surrounding the research that has been conducted and commissioned by the FCC, it is particularly troubling that such research can essentially be immune to re-analysis by interested stakeholders due to the variety of obstacles illustrated here.

Data underlying third-party research

Thus far, this discussion has focused on data transparency and access issues within the context of research conducted or commissioned by government agencies such as the FCC. It is important to recognize that these concerns also extend to the research and data utilized by third parties—such as industry organizations, public interest groups, and scholars—in their participation in regulatory proceedings. Today, any interested stakeholder can submit a study in an FCC proceeding and is under no obligation to make the data utilized in the study available for re-analysis by either the FCC or other stakeholders involved in the proceeding. And, such submissions can prove to be tremendously influential to policymakers' decision-making.

It is important to emphasize that existing safeguards such as the Data Quality Act and the Data Access Act apply only to federally-funded research and data gathering. These safeguards do not apply to the research and data entered into the policymaking process by third parties—though there have been efforts by some stakeholders to argue that even the data underlying third-party research submissions should be publicly accessible (see, e.g., [Center for Regulatory Effectiveness, 2006](#); [Small Market Broadcasters, 2006](#)). At this point, however, the boundaries of the Data Quality Act (as well as the Data Access Act) appear confined to data gathered with federal funds.

A recent controversy surrounding broadcast signal strength data helps illustrate today's rather tangled and unsatisfactory state of affairs in this regard. In 2006, DBS provider EchoStar Satellite argued in court that Section 553 of the Administrative Procedure Act requires that *any* data relied upon by the FCC (or any other federal agency) in decision-making processes must be made available in the public record (*EchoStar Satellite v. FCC, 2006*). In this case, the company sought access to broadcast signal strength data that the FCC used to determine broadcast signal transmission rights under the Satellite Home Viewer Improvement Act of 1999. These data were not submitted in raw form, but rather were analyzed in an engineering report submitted to the FCC by the [National Association of Broadcasters and Association for Maximum Service Television \(2000, February 22\)](#).

The FCC argued that “EchoStar was not entitled to the data because the Commission itself neither had nor relied upon them when it issued its final rule. Rather, the Commission based its analysis upon the *description, methodology, and results* of the study contained in the public comments” (*EchoStar Satellite v. FCC, 2006*, p. 38; emphasis added). This rather tortured logic was not, unfortunately, subjected to scrutiny by the court. The court did not address the issue of EchoStar's right to access the data, on the grounds that EchoStar requested the data only after the Commission had issued its final decision (*EchoStar Satellite v. FCC, 2006*). In any case, a close reading of Section 553 of the [Administrative Procedure Act, 2000](#) provides little support for EchoStar's argument, as the section does not in any meaningful way directly address the issue of access to third-party data that influence regulatory decision-making.

It is important to emphasize that, within this context of the data underlying studies filed by third parties, all of the challenges related to the use and dissemination of commercial data sources (see above) also apply, in addition to those challenges that may arise due to the desires of government agencies or the organizations filing the studies to prevent access to the underlying data. And, the fact that the limited mechanisms that currently are in place to ensure data access and transparency apply primarily to data collected and/or funded by the government means that this entire sector of increasingly influential communications policy research remains largely impenetrable to the kind of open scrutiny and re-analysis that help to preserve the integrity of the policymaking process.

Recommendations

Thus far, this article has illustrated: a) that data transparency and access are fundamental to public policymaking that adheres to basic democratic principles; and b) that, within the context of communications policymaking, minimum standards of data transparency and access are not consistently being met, which is largely due to an incomplete set of legal and regulatory safeguards. This section offers a set of concrete proposals intended to further reduce the gap between the needed levels of data transparency and access and the current reality.

Mandatory data access for all policy research

First, a fundamental legislative requirement needs to be instituted by which *any* study either produced by a government agency or filed

with a government agency by a third party, must make the underlying data used in the study publicly available for re-analysis to some extent. Essentially, this would involve an expansion of the Data Access Act beyond its current focus on data gathered with the support of government grants. Such a requirement would, by making all research findings and conclusions potentially subject to re-analysis, bring the needed level of transparency to this dimension of the policymaking process.

In addition, it would likely have the added effect of curbing the avalanche of “junk science” that is increasingly clogging regulatory and legal proceedings ([Raul & Dwyer, 2003](#)). [Wagner \(2009\)](#) describes the phenomenon of “information capture” that has overtaken the policy-making process in the U.S. According to [Wagner \(2009\)](#), “In the regulatory context, information capture refers to the excessive use of information and related information costs as a means of gaining control over regulatory decision-making” (p. 3). In today's system, stakeholders have powerful “incentives to load as much information as possible into the system” ([Wagner, 2009, p. 3](#)) with little concern about the credibility or reliability of this information. And, of course, the resource imbalances amongst stakeholder groups come to the fore in this context, given that “less-well financed interest groups find it hard to continue to participate” ([Wagner, 2009, p. 3](#)) in this strategic game of attempting to bury policymakers under stacks of studies. The end result, of course, is that objective and credible policy research can get lost amidst the avalanche of the more politically-motivated research conducted to serve the ever growing “market for data” ([Warren, 2002](#)), which is fueled by the existing dynamics of the policymaking process.

Once policy research is conducted in the “sunlight,” so to speak, the likelihood of stakeholders in the process attempting to influence policymakers with blatantly biased or flawed research or data would likely diminish substantially, given that it would become much easier for policymakers and other interested stakeholders to expose such flawed research. With data transparency and access comes an important component of the research checks and balances that are essential to effective and transparent policymaking.

Perhaps the biggest concern that arises from any efforts to increase public accessibility of policy-relevant commercial data sources involves how the various commercial data providers, who currently provide much of the data that support contemporary communications policy research, would respond. Specifically, it seems likely that commercial data providers would resist licensing their data to clients who would be using the data in ways (i.e. for policy research/advocacy) that might subsequently require the data to be made available to a larger audience. Commercial data providers argue that such access would do significant damage to their established business model (see, e.g., [Institute for Public Representation v. Federal Communications Commission, 2008](#)). That is, if some sort of system were put in place to require that commercial data be made available for free to stakeholders in the policy process, the result is lost revenue. This concern becomes compounded by the possibility of the data then circulating further beyond those policy stakeholders who initially obtained them. Under these circumstances, the financial returns for producing such data sets would be diminished, potentially resulting in lower quality data sets, higher initial access costs, or, perhaps even insufficient economic incentives to gather and aggregate data. In this scenario, both the individual data firm and the broader public interest are harmed ([Wolken, 1998](#)).

A number of points need to be raised in response to these concerns. The first is that what we can call here the “policy market” for commercial data providers represents a minuscule component of the overall market that these providers serve. That is, virtually all of the commercial databases that are used in communications policymaking are developed for markets *completely separate* from the fairly narrow context of communications policy analysis being discussed here. Thus, for instance, television, radio, or internet audience data that factor significantly into policy research are produced primarily for a very large, and lucrative, market of advertisers and content providers.

Similarly, communications industry financial data are produced primarily for the finance and investment communities. In fact, the extent to which these commercial data sources are not produced for policy-related purposes, but rather primarily in response to the demands of a completely different marketplace, often mitigates the extent to which these data sources can be effectively utilized for policy analysis (Flamm et al., 2007, November; Lyndon, 1989; Napoli & Karaganis, 2008). Individuals or organizations who purchase such commercial data for policy research purposes represent an utterly tangential and, ultimately, insignificant component of the revenue stream for these commercial data providers.

A related concern arises from the possibility that should commercial data be made more widely available to policy stakeholders, there is an increased likelihood that the data could find their way into the hands of commercial clients, thereby damaging the primary revenue stream and again undermining the production or accessibility of such data. This is a legitimate concern, but one that exists for commercial data providers regardless of the recipient of the data. While it might be a violation of their license agreement, commercial clients can just as easily share commercial databases as any policy researchers. To the extent that the existing protection mechanisms—such as licensing restrictions accompanied by the threat of legal action, and technologically-imposed limitations on file sharing and duplication—currently maintain the integrity of these markets for commercial data, the wider dissemination of such data amongst the relatively small universe of policy researchers should do nothing to alter this situation.

Improved access models and procedures

Moreover, there are, in all likelihood, workable models for maintaining sufficient controls over the access to policy-relevant commercial databases. As has been suggested in other discussions of the data access issue, accessibility to underlying data could potentially be governed by some sort of “need to know” provision that limits access to qualified parties with a demonstrated stake in the decision outcome and limits how the data are used (see e.g., National Research Council, 2005, p. 14).

In order to ensure that the necessary access is extended to the relevant commercial databases, one possible approach that has been proposed is to institute some form of a compulsory licensing scheme for commercial databases. Under such a system, all databases would be registered and deposited with a governmental agency that would enforce database protection and apply any exceptions to database protection based on public interest considerations, such as the necessity of the data for policy research purposes (Lipton, 2003; Wolken, 1998). Such an initiative recognizes the substantial public interest value in many commercial databases, and would, in all likelihood, have negligible effects on the revenue streams of commercial data providers (see below).

In addition, data access can be confined to select locations. Consider, for instance, how access to U.S. General Population Survey data is handled. These data reside in 13 data centers scattered throughout the country. Access to these data centers is highly controlled. Limitations are placed on what researchers can bring into and out of these facilities, in order to protect the privacy of the survey participants. All statistical analyses are done within the confines of the data center, under direct supervision, and the researcher is able to leave only with the aggregate statistical analyses. One participant in the Federal Communications Commission, 2009 Broadband Data Transparency Workshop astutely proposed that such a model be employed for dealing with the confidentiality issues surrounding broadband data—a suggestion that certainly could be applied more broadly to other policy-relevant data sources.⁶

⁶ Statement of Professor Shane Greenstein, Northwestern University, p. 14 of workshop transcript (http://www.ntia.doc.gov/advisory/broadbanddata/Workshop-Transcript_10302009.txt).

The Federal Communications Commission already has taken some important steps in this direction. In 2002, the Commission made much of the underlying data for the studies it commissioned in relation to its media ownership proceeding available online for download and re-analysis (Federal Communications Commission, 2002). The commercial, proprietary status of some of the data used in these studies created complications, however, and some of the data were made available only on a computer terminal physically located at the FCC, with no copying or printing of the data allowed (Federal Communications Commission, 2002). Data access for studies commissioned by the FCC for its 2006 media ownership proceeding were handled in a similar manner (Federal Communications, 2007). However, licensing restrictions associated with Standard and Poor's financial data for radio companies led the Commission to withhold the data set related to one of the FCC's studies on the radio industry (Federal Communications, 2007).

An acceptable threshold of data transparency and access can be achieved via only modest extensions beyond these recent examples. Specifically, if 1) the media ownership data were made available at additional locations around the country; 2) there were no exclusions in terms of available data; and 3) the same access standards and procedures required of the data sets were utilized by all stakeholders filing formal comments/studies in the media ownership proceedings, then the needed levels of data transparency and access necessary would have been achieved.

Realistically assessing the criteria for commercially sensitive information

Another concern that arises in response to any proposals to enhance data access and transparency is that much of the data used in communications policy analysis (whether commercial data sources or data gathered by the federal government) involves commercially sensitive information. These, too, are legitimate concerns (see Anderson & Seltzer, 2009). However, the access controls proposed above would go a long way toward alleviating concerns about sensitive commercial information being obtained and exploited under the guise of the conduct of policy research.

In addition, it is important to emphasize that, in many cases, the industry stakeholders who resist expanding access to commercially sensitive data are, in fact, resisting expanding access to data that often already are available via other sources (see Turner, 2009). Thus, for instance, within the context of broadband deployment data, Frieden (2009) notes that, “Ironically, some carriers that willingly display maps touting their wireless services argue against the FCC providing analogous information about broadband penetration to the public” (p. 104). In other cases, information that regulated firms argue is commercially sensitive can be obtained via licensing commercial databases. This kind of data availability needs to be taken into consideration when claims of confidentiality are assessed. If the data are obtainable by other available means, then any claim that the data represent sensitive or confidential commercial information is essentially unfounded.

The specific rationales put forth that certain data should be treated as confidential also need to be taken into consideration. Returning to the example of broadband deployment data, the argument put forth that granular deployment data should be treated as confidential is often premised on the notion that telecommunications service providers could utilize the data to identify regions in which to try to compete with incumbent providers (Frieden, 2009). Thus, essentially, the availability of the data would facilitate enhanced competition in the provision of broadband services. This is of course something that is desired by policymakers and certainly serves the public interest in this era when broadband deployment is seen as fundamental to economic development (Frieden, 2009). Even the FCC has stated that “wide dissemination of the information promotes a more informed more efficient market” (Federal Communications Commission, 2000, p.

7757). Moreover, as Frieden (2009) notes, “Identifying locations where a specific carrier does not operate provides little, if any, potential financial harm to a carrier that already has determined that service there does not make financial sense” (p. 119). This perspective certainly weakens the incentive for policymakers to treat such data as confidential; and, when combined with the fact described above, that competitors often can obtain commercially sensitive information via other means, creates a set of circumstances that undermines the logic of restricting access to such data for policy research.

Excluding data from the deliberative process FOIA exemption

As was discussed above, there seems to be a fundamental disconnect between the underlying rationales of the deliberative process FOIA exemption and the application of such an exemption to the factual data that underlie the policy analyses relied upon by federal agencies. For this reason, data should be explicitly excluded from the deliberative process exemption (Jensen, 1999). Such an exclusion in no way weakens the extent to which the exemption can serve its primary functions of allowing for the free flow of ideas within government agencies and protecting against policy deliberations prematurely reaching the citizenry and being misinterpreted as policy actions. The notion that factual data can be withheld from the public on behalf of protecting the confidentiality of the deliberative process within government agencies not only fundamentally overstates the level of confidentiality that should be conferred upon such deliberations, but also undermines one of the primary tools (data access and re-analysis) by which government agency officials can be held accountable.

Federal Advisory Committee on Data Quality, Integrity and Access

Finally, given the complex and evolving nature of the issues surrounding data access and usage in communications policymaking, and given the range of initiatives proposed thus far, a Federal Advisory Committee on Data Quality, Integrity, and Access should be created. Such a committee would be comprised of FCC and NTIA personnel, industry representatives, academic researchers, and members of the public interest and advocacy communities. Such a committee could play an integral role in constructing and overseeing the system of archiving and granting researcher access to data sources sketched out above.

In addition, this committee could serve in an advisory capacity to policymakers on a number of fronts, including the establishment of specific baseline standards for policymakers' data needs; assessing the quality and integrity of the various data sets relied upon not only by the relevant government agencies, but also by the various external stakeholders that submit research in individual proceedings; and in terms of conducting regular, systematic inventories and assessments of the various data-gathering forms that agencies such as the FCC and NTIA employ to obtain data from the organizations under their regulatory authority. Such a committee could also play a role in these agencies' responses to Data Quality Act complaints, as well as in the administration of the external peer review process that is now required for government-sponsored research.

Conclusion

This article has used U.S. communications policymaking as a case study of the issue of public availability of the data that inform policymaking. The fundamental premise of this analysis has been that public policy should be made with publicly available data. A truly open government depends upon adherence to such a principle.

While this analysis has focused on communications policymaking, it is important to emphasize that these concerns have permeated other policy areas as well, such as health, economic, and environ-

mental policy (see, e.g., Anderson & Seltzer, 2009; Cohen & Hahn, 1999; Wagner, 2009; Wagner & Michaels, 2004). In all of these contexts, the same issues of resource imbalances among policy stakeholders, accessibility of commercial data sources, and overbroad protections for commercially sensitive information and government agencies' deliberative processes, are present to varying degrees.

Many of the proposals put forth in the previous section are applicable to policy-relevant data more broadly, not just within the context of communications policy. Moving forward, policymakers need to consider—in a way that spans the full range of policy contexts and relevant government agencies—legislative initiatives that: a) acknowledge the public policy relevance of many commercial databases and the need for some form of controlled public access to such databases; b) clarify, and to some extent constrain, the existing exceptions to providing public access to policy-relevant data; and c) integrate the largely neglected realm of privately-funded research utilized in policymaking into existing data access safeguards. Evidence-driven approaches to policymaking have taken hold across virtually all policymaking contexts, making it essential that the fundamentally democratic principles of data transparency and access are upheld.

References

- Administrative Procedure Act (2000). 5 U.S.C. § 553.
- Anderson, M., & Seltzer, W. (2009). Federal statistical confidentiality and business data: Twentieth century challenges and continuing issues. *Journal of Privacy and Confidentiality*, 1(1), 7–52.
- Broadband Data Improvement Act (2008). P.L. 110-385, 47 USC. 1301.
- Burgee, K. S. (2006, September 26). Letter from Kirk S. Burgee, Associate Bureau Chief, Wireline Competition Bureau, FCC, to Drew Clark, Sr. Fellow and Project Mgr., Center for Public Integrity Retrieved January 23, 2009, from <http://www.publicintegrity.org/docs/FCCReplytoFOIA.pdf>.
- Center for Public Integrity v. FCC, 515 F.Supp.2d 167 (D.D.C. 2007)
- Quadrennial Review—Review of the commissioner broadcast ownership rules and other rules adopted pursuant to section 202 of the Telecommunications Act of 1996, Formal Comments (accessible via FCC Electronic Comment Filing System)
- Coglianesi, C., Zeckhauser, R., & Parson, E. (2004). Seeking truth for power: Informational strategy and regulatory policymaking. *Minnesota Law Review*, 89, 277–341.
- Cohen, L., & Hahn, R. W. (1999). A solution to concerns over public access to scientific data. *Science*, 285, 535–536.
- Cramer, B. W. (2009). “The nation’s broadband success story”: The secrecy of FCC broadband infrastructure statistics. *Hastings Communications and Entertainment Law Journal*, 31, 339–368.
- Data Access Act of 1998, Pub. L. No. 105-277, 112 Stat. 2681.
- Data Quality Act of 2000, § 515, Pub. L. No. 106-554, 114 Stat. 2763.
- Duwadi, K., Roberts, S., & Wise, A. (2007). Ownership structure and robustness of media: FCC media ownership study #2 Retrieved January 23, 2010, from http://fjallfoss.fcc.gov/edocs_public/attachmatch/DA-07-3470A3.pdf.
- EchoStar Satellite v. FCC, 457 F.3d 31, 39 (D.C. Cir. 2006)
- Federal Communications Commission. (2000). Local competition and broadband reporting, report and order, 15 F.C.C.R. 7717.
- Federal Communications Commission. (2002). 2002 Biennial Regulatory Review—Review of the Commission’s broadcast ownership rules and other rules adopted pursuant to section 202 of the Telecommunications Act of 1996, Protective Order, 17 F.C.C.R. 22,178.
- Federal Communications Commission (2004). Broadcast Localism, Notice of Inquiry, 19 F.C.C.R. 12, 425.
- Federal Communications Commission (2007). 2006 Quadrennial Review—Review of the Commission’s Broadcast Ownership Rules and Other Rules Adopted Pursuant to Section 202 of the Telecommunications Act of 1996, Protective Order. Retrieved January 28, 2010, from http://fjallfoss.fcc.gov/edocs_public/attachmatch/DA-07-3741A1.pdf
- Federal Communications Commission. (2009). A national broadband plan for our future, protective order Retrieved January 28, 2010, from http://hraunfoss.fcc.gov/edocs_public/attachmatch/DA-09-2187A1_Rcd.pdf.
- Federal Communications Commission (2010). Comment Sought on Free Press Request to Review Form 477 Data and Request for Protective Order. WC Docket No. 10-75.
- Flamm, K., Friedlander, A., Horrigan, J., & Lehr, W. (2007, November). Measuring broadband: Improving communications policymaking through better data collection. Report of a workshop co-sponsored by the Pew Internet & American Life Project. : The University of Texas at Austin, and the Massachusetts Institute of Technology.
- Ford, G. S. (2007, May 17). Written statement of George S. Ford, Chief Economist, Phoenix Center for Advanced Legal & Economic Public Policy Studies, before the House of Representatives Committee on Energy and Commerce, Subcommittee on Telecommunications and the Internet Retrieved January 23, 2010, from http://archives.energycommerce.house.gov/cmtc_mtgs/110-ti-hrg.051707.Ford-testimony.pdf.

- Frieden, R. (2009). Lies, damn lies and statistics: Developing a clearer assessment of market penetration and broadband competition in the United States. *Virginia Journal of Law & Technology*, 14(2), 100–125.
- Heinrich, C. J. (2007). Evidence-based policy and performance management: Challenges and prospects in two parallel movements. *American Review of Public Administration*, 37, 255–277.
- Institute for Public Representation v. Federal Communications Commission (2007). Complaint, No. 07CV02092 (D.D.C. 2007), dismissed, 2007 WL 2900431 (D.D.C. July 9, 2008).
- Institute for Public Representation v. Federal Communications Commission (2008). Intervener defendant BIA Financial Network Inc's memorandum of points and authorities in support of its motion for summary judgment. : United States District Court for the District of Columbia.
- Jensen, K. D. (1999). The reasonable government official test: A proposal for the treatment of factual information under the federal deliberative process privilege. *Duke Law Journal*, 49(2), 561–599.
- Knight Commission on the information needs of communities in a democracy (2009). *Informing communities: sustaining democracy in the digital age*. Washington, DC: Aspen Institute.
- Lipton, J. (2003). Balancing private rights and public policies: Reconceptualizing property in databases. *Berkeley Technology Law Journal*, 18, 774–852.
- Lyndon, M. L. (1989). Information economics and chemical toxicity: Designing laws to produce and use data. *Michigan Law Review*, 87(7), 1795–1861.
- Napoli, P. M. (2008). The paradoxes of media policy analysis: Implications for public interest media regulation. *Administrative Law Review*, 60(4), 801–812.
- Napoli, P. M., & Karaganis, J. (2008). Toward a federal data agenda for communications policymaking. *CommLaw Conspectus*, 16(1), 53–96.
- Napoli, P. M., & Seaton, M. (2007). Necessary knowledge for communications policymaking: Information asymmetries and commercial data access and usage in the policymaking process. *Federal Communications Law Journal*, 59(2), 295–330.
- National Association of Broadcasters and Association for Maximum Service Television. (2000, February 22). *Establishment of an improved model for predicting the broadcast television field strength received at individual locations* Engineering Statement in Support of Comments, ET Docket No. 00-11.
- National Parks and Conservation Association v. Morton, 498 F.2d 765 (D.C. Cir. 1974)
- National Research Council (2005). Expanding access to research data: Reconciling risks and opportunities. Washington, DC: National Academies Press.
- Nelkin, D. (1979). Scientific knowledge, public policy, and democracy: A review essay. *Science Communication*, 1, 106–122.
- Open government: A progress report to the American people (2009, December). Retrieved January 23, 2010, from <http://www.whitehouse.gov/open>.
- Perko, M.S. (2007, January 4). Letter from Michael S. Perko, Chief, Office of Communications & Industry Information, FCC, to Marvin Ammori, Institute for Public Representation, Georgetown University Law Center (on file with authors).
- Raul, A. C., & Dwyer, J. W. (2003). Science in the regulatory process: "Regulatory Daubert": A proposal to enhance judicial review of agency science by incorporating Daubert principles into administrative law. *Law and Contemporary Problems*, 66, 7–44.
- Robinson, D., Yu, H., Zeller, W. P., & Felton, E. W. (2008/2009). Government data and the invisible hand. *Yale Journal on Law & Technology*, 11, 160–175.
- Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public Administration*, 80(1), 1–22.
- Sandoval, C.J.K. (in press). Minority commercial radio ownership in 2009: FCC licensing and consolidation policies, entry windows, and the nexus between ownership, diversity, and service in the public interest. In P.M. Napoli & M. Aslama (Eds.), *Communications research in action: Scholar-activist collaborations for a democratic public sphere*. New York: Fordham University Press
- Small Market Broadcasters Coalition (2006, December 7). 2006 Quadrennial Review—Review of the Commissioner Broadcast Ownership Rules and Other Rules Adopted Pursuant to Section 202 of the Telecommunications Act of 1996, Coalition Request for Underlying Data (accessible via FCC Electronic Comment Filing System).
- Turner, D. (2009, September 4). *In the matter of inquiry concerning the deployment of advanced telecommunications capability to all Americans in a reasonable and timely fashion, and possible steps to accelerate such deployment pursuant to Section 706 of the Telecommunications Act of 1996, as amended by the Broadband Data Improvement Act* Comments of Free Press. Retrieved January 25, 2010, from <http://fjallfoss.fcc.gov/ecfs/document/view?id=7020037663>.
- U.S. Government Accountability Office. (2006). *Broadband deployment is extensive throughout the United States, but it is difficult to assess the extent of deployment gaps in rural areas* Retrieved January 23, 2010, from <http://www.gao.gov/new.items/d06426.pdf>.
- U.S. Government Accountability Office. (2009). *Current broadband measures have limitations, and new measures are promising but need improvement* Retrieved January 23, 2010, from <http://www.gao.gov/new.items/d1049.pdf>.
- U.S. House of Representatives (2007, May 17). *Committee on Energy and Commerce, hearing on broadband mapping* Retrieved November 16, 2007, from http://energycommerce.house.gov/cmtc_mtgs/110-ti-hrg.051707.Broadband.mapping.shtml.
- Wagner, W.E. (2009). Administrative law, filter failure, and information capture. Unpublished working paper. Retrieved January 14, 2010, from http://works.bepress.com/wendy_wagner/1.
- Wagner, W., & Michaels, D. (2004). Equal treatment for regulatory science: Extending the controls governing the quality of public research to private research. *American Journal of Law & Medicine*, 30, 119–154.
- Warren, E. (2002). The market for data: The changing role of social sciences in shaping the law. *Wisconsin Law Review*, 2002, 1–43.
- Wolken, J. C. (1998). Just the facts, ma'am: A case for uniform federal regulation of information databases in the new information age. *Syracuse Law Review*, 48, 1263–1305.

Philip M. Napoli is a Professor in the Graduate School of Business and Director of the Donald McGannon Communication Research Center at Fordham University. His research interests focus on media institutions and media policy. He is the author of the books *Foundations of Communications Policy: Principles and Process in the Regulation of Electronic Media* (Hampton Press, 2001); *Audience Economics: Media Institutions and the Audience Marketplace* (Columbia University Press, 2003) and *Audience Evolution: New Technologies and the Transformation of Media Audiences* (Columbia University Press, 2010).

Joe Karaganis is the Media, Technology, and Culture Program Director at the Social Science Research Council in New York. His research focuses on the relationship between digital convergence and cultural production. His books include *Structures of Participation in Digital Culture* (Social Science Research Council, 2008) and *Toward a Detente in Media Piracy* (forthcoming). He has also worked as a consultant for the Ford and Rockefeller Foundations.